Estimation of spectral bounds in gradient algorithms

Luc Pronzato, Anatoly A. Zhigljavsky, Elena Bukina

Estimation of spectral bounds in gradient algorithms*

L. PRONZATO[1,2], A. ZHIGLJAVSKY[3] and E. BUKINA[2]
[1] Author for correspondence
[2] Laboratoire I3S, CNRS/Université de Nice-Sophia Antipolis
Bât. Euclide, Les Algorithmes, 2000 route des lucioles, BP 121
06903 Sophia Antipolis cedex, France
[3] School of Mathematics, Cardiff University
Senghennydd Road, Cardiff, CF24 4YH, UK
pronzato@i3s.unice.fr
ZhigljavskyAA@cf.ac.uk
bukina@i3s.unice.fr

December 14, 2012

Abstract

We consider the solution of linear systems of equations $Ax = b$, with A a symmetric positive-definite matrix in $\mathbb{R}^{n\times n}$, through Richardson-type iterations or, equivalently, the minimization of convex quadratic functions $(1/2)(Ax, x) - (b, x)$ with a gradient algorithm. The use of step-sizes asymptotically distributed with the arcsine distribution on the spectrum of A then yields an asymptotic rate of convergence after $k < n$ iterations, $k \to \infty$, that coincides with that of the conjugate-gradient algorithm in the worst case. However, the spectral bounds m and M are generally unknown and thus need to be estimated to allow the construction of simple and cost-effective gradient algorithms with fast convergence. It is the purpose of this paper to analyse the properties of estimators of m and M based on moments of probability measures ν_k defined on the spectrum of A and generated by the algorithm on its way towards the optimal solution. A precise analysis of the behavior of the rate of convergence of the algorithm is also given. Two situations are considered: (i) the sequence of step-sizes corresponds to i.i.d. random variables, (ii) they are generated through a dynamical system (fractional parts of the golden ratio) producing a low-discrepancy sequence. In the first case, properties of random walk can be used to prove the convergence of simple spectral bound estimators based on the first moment of ν_k. The second option requires a more careful choice of spectral bounds estimators but is shown to produce much less fluctuations for the rate of convergence of the algorithm.

keywords estimation of leading eigenvalues; arcsine distribution; gradient algorithms; conjugate gradient; Fibonacci numbers

MSC 65F10; 65F15

*Part of this work was accomplished while the first two authors were invited at the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK; the support of the INI and of CNRS is gratefully acknowledged. The work of E. Bukina was partially supported by the EU through a Marie-Curie Fellowship (EST-SIGNAL program: http://est-signal.i3s.unice.fr) under the contract Nb. MEST-CT-2005-021175.

1 Introduction and motivation

For $\{\nu_k\}_{k=0}^{\infty}$ a sequence of probability measures supported on a real interval $[m, M]$, the sequence of first moments $\mu_1^{(k)} = \int_m^M t\,\nu_k(dt)$ gives obvious estimators of m and M through

$$\widehat{m}_k = \min_{j=0,\ldots,k} \mu_1^{(j)} \quad \text{and} \quad \widehat{M}_k = \max_{j=0,\ldots,k} \mu_1^{(j)}\,. \tag{1}$$

We consider the behavior of estimators (1) and their extensions defined below when ν_k are probability measures associated with a gradient algorithm for the minimization of a convex quadratic function with matrix $A \in \mathbb{R}^{n\times n}$ (symmetric positive-definite) and $\mu_1^{(k)} = (Ag_k, g_k)/(g_k, g_k)$, with g_k, the gradient at step k, obeying the recurrence equations

$$g_{k+1} = g_k - \gamma_k A g_k\,,\ k = 0, 1, 2\ldots \tag{2}$$

Here, $\gamma_k > 0$ is the step-size at iteration k and is determined by some rule that characterizes the algorithm. For instance, $\gamma_k = 1/\mu_1^{(k)}$ for the Steepest Descent (SD) algorithm, $\gamma_k = (Ag_k, g_k)/(A^2 g_k, g_k)$ for the method of Minimum Residues (MR), see [8], [18, p. 134], and $\gamma_k = 1/\mu_1^{(k-1)}$ for the method of Barzilai and Borwein [1]. The algorithms considered in [3, 12, 20] rely on the generation of an infinite sequence of step-sizes γ_k (possibly random), such that $\beta_k = 1/\gamma_k \in [m, M]$ for all k, with m and M respectively the minimum and maximum eigenvalues of A. As shown in [16], when the sequence $\{\beta_k\}$ is asymptotically distributed with the arcsine density in $[m, M]$, then the asymptotic rate of convergence is competitive compared to that of Conjugate Gradients (CG) [7], Conjugate Residuals (CR) [5, p.547], or other methods based on Krylov spaces, like *e.g.* MINRES [11] (it coincides with that exhibited by CG and CR in the worst case, in terms of choice of the starting point and locations of the $n-2$ internal eigenvalues of A in (m, M), when the algorithm is stopped before n iterations). Such algorithms, with step-sizes generated externally, are simpler than CR and CG and thus of particular interest in situations where n is so large that the algorithm is stopped well before n iterations (in particular, it is shown in [22] that for some sequences of step-sizes the number of scalar products to be computed out of k iterations only grows as $\mathcal{O}(\log k)$, see also Sect. 3.3 and 5). The generation of suitable sequences of step-sizes γ_k requires, however, the knowledge of the spectral bounds m and M. Since they are usually unknown, it is suggested in [22, 16] to estimate them through the evaluation of moments of probability measures generated by the algorithm itself. It is the purpose of this paper to analyse the asymptotic properties of the estimators (1) of the leading eigenvalues of A. In particular, two situations will be considered: (i) the γ_k form a sequence of i.i.d. random variables, (ii) they are constructed from a low discrepancy sequence. In addition, a more precise analysis of the behavior of the rate of convergence of the algorithm than in [22, 16] will be provided (Sect. 4).

2 A sequence of probability measures associated with a gradient algorithm

Consider a linear system of equations

$$Ax = b\,, \tag{3}$$

where $x \in \mathbb{R}^n$ is an unknown vector, A is a $n \times n$ symmetric positive-definite matrix such that $0 < m = \inf_{(z,z)=1}(Az, z) < M = \sup_{(z,z)=1}(Az, z) < \infty$ and b is a given vector in $\mathbb{R}^n$.

Equivalently, one may consider the minimization of the convex quadratic function

$$f(x) = \frac{1}{2}(Ax, x) - (b, x)\,. \tag{4}$$

The gradient $g_k = Ax_k - b$ then corresponds to the (minus) residual of the system (3) at x_k. Richardson-like methods for solving (3) correspond to gradient algorithms for minimizing (4) and obey to the following iterations

$$x_{k+1} = x_k - \gamma_k g_k\,,\ k = 0, 1, 2 \dots\,, \tag{5}$$

where $x_0 \in \mathbb{R}^n$ is a starting vector and $\gamma_k > 0$ is the step-size at iteration k. The methods to be considered are of particular interest in situations where the dimension n can be very large (it could even be infinite, with A a self-adjoint operator in a Hilbert space, but we shall restrict the presentation to the finite dimensional situation) and we shall assume that n is much larger than the number of iterations k_* needed to achieve the precision required (formally, $k_* = o(n)$ with $k_* \to \infty$ in asymptotic considerations).

The iterations (5) can be rewritten in terms of gradients g_k, which gives (2), with $g_0 = Ax_0 - b \in \mathbb{R}^n$ the initial gradient. We define the rate of convergence of the algorithm (5) towards the solution at iteration j as

$$r_j = \frac{(g_{j+1}, g_{j+1})}{(g_j, g_j)}\,;$$

the rate of convergence after k iterations is then

$$R_k = \prod_{j=0}^{k-1} r_j = \frac{(g_k, g_k)}{(g_0, g_0)}\,. \tag{6}$$

Other convergence rates, which are asymptotically equivalent to r_j (see [14, Th. 6]), can also considered. In particular,

$$r'_j = \frac{f(x_{j+1}) - f(x^*)}{f(x_j) - f(x^*)} = \frac{(A^{-1}g_{j+1}, g_{j+1})}{(A^{-1}g_j, g_j)}$$

is often used when minimizing a quadratic function (4), with $x^* = A^{-1}b$ its minimizer.

The method of Steepest-Descent (SD) chooses γ_k in (5) that minimizes r'_k and the method of Minimum Residues (MR) chooses γ_k that minimizes r_k, both are myopic and only look one-step forward. The method of Conjugate Gradients (CG) minimizes

$$R'_k = \prod_{j=0}^{k-1} r'_j = \frac{f(x_k) - f(x^*)}{f(x_0) - f(x^*)} = \frac{(A^{-1}g_k, g_k)}{(A^{-1}g_0, g_0)}$$

with respect to the sequence $\gamma_0, \gamma_1, \dots, \gamma_{k-1}$ and the method of Conjugate Residuals (CR) does the same with R_k; we shall denote $R_k^{CR} = \min_{\gamma_0,\dots,\gamma_{k-1}} R_k$. Although CR minimizes R_k for all k, for any $k < n$ one may nevertheless have, in the worst case with respect to the starting point x_0 and eigenvalues $\lambda_2, \dots, \lambda_{n-1}$, $R_k^{CR} = R_k^*$, where

$$R_k^* = \left(\frac{R_\infty^{k/2} + R_\infty^{-k/2}}{2}\right)^{-2} = C_k^{-2}\left(\frac{\rho+1}{\rho-1}\right),$$

with $\rho = M/m$ the condition number of A, $C_k(\cdot)$ the k-th Chebyshev polynomial of the first kind, $C_k(t) = \cos[k \arccos(t)] = (1/2)[(t + \sqrt{t^2-1})^k + (t - \sqrt{t^2-1})^k]$, and

$$R_\infty = \lim_{k\to\infty} (R_k^*)^{1/k} = \left(\frac{\sqrt{\rho}-1}{\sqrt{\rho}+1}\right)^2 , \tag{7}$$

see [4, 15]. Hence, although one regularly observes values of R_k^{CR} that are significantly smaller than R_k^* for $k < n$ (and although $R_n^{CR} = 0$, that is, the solution is found exactly in n iterations in the exact arithmetic), for any $k < n$ one has $\max_{x_0,A} R_k^{CR} = R_k^*$, with $(R_k^*)^{1/k}$ decreasing monotonically to R_∞ as $k \to \infty$. As shown in [16], the same asymptotic rate R_∞ can be obtained when the sequence $\{\beta_k\} = \{1/\gamma_k\}$ is generated externally with some suitable distribution in $[m, M]$; see also Sect. 4.

From (2), the gradient g_k after k iterations can be written as

$$g_k = P_k(A) g_0 \,, \tag{8}$$

where P_k denotes the polynomial $P_k(A) = (I - \gamma_{k-1}A)(I - \gamma_{k-2}A) \ldots (I - \gamma_0 A)$.

Let $m = \lambda_1 \le \ldots \le \lambda_n = M$ be the eigenvalues of A and $\{q_1, \ldots, q_n\}$ be the set of corresponding orthonormal eigenvectors (we assume that no information is available about the eigenvalues λ_i and eigenvectors q_i, $i = 1, \ldots, n$, and that the condition number M/m may be large). When decomposing the initial vector g_0 in the basis $\{q_1, \ldots, q_n\}$ as $g_0 = \sum_{i=1}^n \alpha_i q_i$, (8) implies

$$g_k = \sum_{i=1}^n \alpha_i P_k(\lambda_i) q_i \tag{9}$$

for all $k \ge 1$. The squared L_2-norm of g_0 is $\|g_0\|^2 = (g_0, g_0) = \sum_{i=1}^n \alpha_i^2$ and the squared L_2-norm of g_k is thus $\|g_k\|^2 = (g_k, g_k) = \sum_{i=1}^n \alpha_i^2 P_k^2(\lambda_i)$. The convergence rate (6) after k iterations is then given by

$$R_k = \frac{\sum\limits_{i=1}^n \alpha_i^2 P_k^2(\lambda_i)}{\sum\limits_{i=1}^n \alpha_i^2} = \sum_{i=1}^n p_i^{(0)} P_k^2(\lambda_i) \,,$$

where $p_i^{(0)} = \alpha_i^2 / \sum\limits_{j=1}^n \alpha_j^2 \ge 0$ and $\sum_{i=1}^n p_i^{(0)} = 1$.

Without loss of generality all α_i^2 can be assumed to be strictly positive. Indeed, if $\alpha_i = 0$ for some i then the matrix $A = \sum_{j=1}^n \lambda_j q_j q_j^\top$ can be replaced with $\tilde{A} = \sum_{j\ne i} \lambda_j q_j q_j^\top$; the equality $\alpha_i = 0$ would mean that $(A\, x_0, q_i) = (b, q_i)$ (and therefore $(A\, x_k, q_i) = (b, q_i)$ for all k).

From (9), $(g_k, q_i) = \alpha_i P_k(\lambda_i)$. We define

$$p_i^{(k)} = \frac{(g_k, q_i)^2}{(g_k, g_k)} = \frac{\alpha_i^2 P_k^2(\lambda_i)}{\sum_{j=1}^n \alpha_j^2 P_k^2(\lambda_j)}$$

and interpret this as a mass at λ_i. Then, the measure ν_k defined by its masses $\nu_k(\lambda_i) = p_i^{(k)}$ at $\lambda = \lambda_i$ $(i = 1, \ldots, n)$ characterizes the normalized vector $g_k/\|g_k\|$ (up to the signs of the (g_k, q_i), which are irrelevant for analyzing the behavior of the algorithm). For any real α, define $\mu_\alpha^{(k)}$ as the α-th moment of the probability measure ν_k:

$$\mu_\alpha^{(k)} = \mu_\alpha(\nu_k) = \sum_{i=1}^n \lambda_i^\alpha p_i^{(k)} = \frac{(A^\alpha g_k, g_k)}{(g_k, g_k)} \,. \tag{10}$$

Using the basic iteration (2), we obtain the following updating formula which expresses the measure ν_{k+1} through the measure ν_k:

$$p_i^{(k+1)} = \nu_{k+1}(\lambda_i) = \frac{\alpha_i^2 P_{k+1}^2(\lambda_i)}{(g_{k+1}, g_{k+1})} = \frac{\alpha_i^2 (1-\gamma_k\lambda_i)^2 P_k^2(\lambda_i)}{(g_{k+1}, g_{k+1})} = \frac{(1-\gamma_k\lambda_i)^2 p_i^{(k)}}{r_k}\,,\ i=1,\dots,n\,, \quad (11)$$

where $k \geq 0$ and

$$r_k = \frac{(g_{k+1}, g_{k+1})}{(g_k, g_k)} = \frac{(g_k, g_k) - 2\gamma_k(Ag_k, g_k) + \gamma_k^2(A^2 g_k, g_k)}{(g_k, g_k)} = 1 - 2\gamma_k\mu_1^{(k)} + \gamma_k^2\mu_2^{(k)}\,. \quad (12)$$

3 Estimation of the leading eigenvalues of A

3.1 Defining the estimators

Take any probability measure ν on $[m, M]$ with $0 < m < M < \infty$ and denote by μ_α its moment of order α, $\mu_\alpha = \mu_\alpha(\nu) = \int_m^M t^\alpha\,\nu(dt)$ (so that $\mu_\alpha(\nu_k)$ is defined by (10)). The Cauchy-Schwarz inequality implies $\mu_{\alpha+2}\mu_\alpha \geq (\mu_{\alpha+1})^2$ for any α. Moreover, $t(M-t) \geq 0$ for all $t \in [m, M]$ so that $\int_m^M t^\alpha(M-t)\,\nu(dt) = M\mu_\alpha - \mu_{\alpha+1} \geq 0$; that is, $\mu_{\alpha+1}/\mu_\alpha \leq M$. Similarly, $m \leq \mu_{\alpha+1}/\mu_\alpha$. We thus obtain the following chain of inequalities

$$m \leq \mu_1^{(k)} \leq \frac{\mu_2^{(k)}}{\mu_1^{(k)}} \leq \frac{\mu_3^{(k)}}{\mu_2^{(k)}} \leq \frac{\mu_4^{(k)}}{\mu_3^{(k)}} \leq \cdots \leq M\,, \quad (13)$$

which are valid for all $k = 0, 1, \dots$; note that $\mu_0^{(k)} = 1$ for all k.

In what follows we shall restrict our attention to the estimators of m and M defined by

$$\widehat{m}_k = \min_{j=0,\dots,k} \mu_1^{(j)}\,,\ \widehat{M}_k^{(i)} = \max_{j=0,\dots,k} \mu_i^{(j)}/\mu_{i-1}^{(j)}\,,\ i \geq 1\,. \quad (14)$$

According to (13), the larger i in $\widehat{M}_k^{(i)}$ the more precise the estimation of M, which has a significant influence on the behavior of the algorithm, see [22]. Calculating high order moments has some computational cost, however, and a compromise must be made. The algorithm presented in Sect. 5 uses $i = 4$.

Let $\{\alpha_k\}$ denote a sequence in $[-1, 1]$ with asymptotic distribution function $F_\alpha(\cdot)$ symmetric with respect to zero (different types of sequences will be considered below). We shall consider algorithms defined as follows: we initiate (5) by two SD iterations with $\gamma_k = 1/\mu_1^{(k)}$, or two MR iterations with $\gamma_k = \mu_1^{(k)}/\mu_2^{(k)}$; for each subsequent iteration the inverse step-size $\beta_k = 1/\gamma_k$ is obtained by rescaling the k-th element α_k of the sequence $\{\alpha_k\}$ into $[\widehat{m}_k, \widehat{M}_k]$, that is,

$$\beta_k = \alpha_k(\widehat{M}_k - \widehat{m}_k)/2 + (\widehat{M}_k + \widehat{m}_k)/2\,. \quad (15)$$

The assumption that the β_k are generated by symmetric pairs in $[m, M]$, that is, $\beta_{2j+1} = M + m - \beta_{2j}$, is used in [16] to derive the expression for the asymptotic convergence rate R of the algorithm, $R = \lim_{k\to\infty} R_k^{1/k}$; see also Sect. 4. This is why we shall also consider the case when (15) is replaced by

$$\begin{cases} \beta_{2j} = \alpha_j(\widehat{M}_{2j} - \widehat{m}_{2j})/2 + (\widehat{M}_{2j} + \widehat{m}_{2j})/2\,, \\ \beta_{2j+1} = -\alpha_j(\widehat{M}_{2j+1} - \widehat{m}_{2j+1})/2 + (\widehat{M}_{2j+1} + \widehat{m}_{2j+1})/2\,. \end{cases} \quad (16)$$

As shown in [22], using the largest β first in a pair $(\beta_{2j}, \beta_{2j+1})$ permits to improve the monotonicity of the algorithm (5). We shall thus also consider the case when

$$\left\{ \begin{array}{l} \beta_{2j} = |\alpha_j|(\widehat{M}_{2j} - \widehat{m}_{2j})/2 + (\widehat{M}_{2j} + \widehat{m}_{2j})/2\,, \\ \beta_{2j+1} = -|\alpha_j|(\widehat{M}_{2j+1} - \widehat{m}_{2j+1})/2 + (\widehat{M}_{2j+1} + \widehat{m}_{2j+1})/2\,, \end{array} \right. \tag{17}$$

Lemma 1 below shows that $(\widehat{m}_k + \widehat{M}_k)/2$ converges to $(m+M)/2$ so that the symmetry condition with respect to $(m+M)/2$ will be asymptotically satisfied.

3.2 Consistency of $\widehat{m}_k$ and $\widehat{M}_k$

The estimators $\widehat{m}_k$ and $\widehat{M}_k$ satisfy the following asymptotic symmetry property.

Lemma 1 *Assume that in algorithm* (5) $\beta_k = 1/\gamma_k$ *is generated according to one of the rules* (15), (16) *or* (17), *with* $\widehat{m}_k$ *and* $\widehat{M}_k = \widehat{M}_k^{(i)}$ *given by* (14) *for all* k *for some* $i \geq 1$ *and* $\{\alpha_k\}$ *having an asymptotic distribution function* $F_\alpha(\cdot)$ *in* $[-1, 1]$ *symmetric with respect to zero. Then we have*

$$M - M_\infty = m_\infty - m \geq 0\,, \tag{18}$$

where $m_\infty = \lim_{k\to\infty} \widehat{m}_k$ *and* $M_\infty = \lim_{k\to\infty} \widehat{M}_k$.

The proof, based on establishing a contradiction if we assume that the symmetry condition (18) is violated, is given in Sect. 6.

To obtain a more precise characterization of the limiting behaviors of $\widehat{m}_k$ and $\widehat{M}_k$ we shall make use of the following property, shown in [16].

Theorem 1 *Set* $\beta_k = 1/\gamma_k$ $(k = 0, 1, \ldots)$ *and assume that* $\beta_k > 0$ *and* $\beta_k \notin \{m, M\}$ *for all* k *and that the sequence* $\{\beta_k\}$ *has an asymptotic distribution function* $F_\beta(\cdot)$ *which is supported on an interval* $[m', M']$ *with* $0 < m' \leq M' < \infty$. *Suppose, moreover, that this limiting distribution satisfies*

$$\int \log(t-\lambda)^2\, dF_\beta(t) < \max\left\{ \int \log(M-t)^2\, dF_\beta(t)\,,\ \int \log(t-m)^2\, dF_\beta(t) \right\}, \tag{19}$$

for all $\lambda \in (m, M)$. *Then the algorithm* (5) *associated with the sequence* $\{\beta_k\}$ *is such that* $\lim_{k\to\infty} \nu_k(\lambda_i) = 0$ *for all* $i = 2, \ldots, n-1$. *Furthermore, there exist constants* $C > 0, k_0 > 0$ *and* $0 \leq \theta < 1$ *such that* $\sum_{i=2}^{n-1} \nu_k(\lambda_i) \leq C\theta^k$ *for* $k > k_0$.

The main condition in Th. 1 is (19); it implies that the ratio $P_k^2(\lambda)/(P_k^2(m) + P_k^2(M))$ tends to 0 (as $k \to \infty$) exponentially fast for any $\lambda \in (m, M)$. It means that once the attraction of the sequence $\{\nu_k\}$ to the set of measures supported at m and M is obtained, *i.e.* $\nu_k(m) + \nu_k(M) \to 1$, it is roughly enough to consider the behavior obtained for two-point measures.

Remark 1 *The attraction of the sequence* $\{\nu_k\}$ *to the set of measures supported at* m *and* M *does not imply that* $\widehat{m}_k \to m$ *and* $\widehat{M}_k \to M$. *Indeed, consider the case when the sequence of step-sizes is self-generated by the algorithm itself. For instance, for SD we have* $\beta_k = \mu_1^{(k)}$ *for all* k *and the limiting measure for* $\{\beta_k\}$ *is the two-point measure allocating weights 1/2 at* z *and* $M + m - z$ *for some* $z \in (m, M)$. *The condition* (19) *is then equivalent to* z *belonging to the stability interval defined in [13, 14], see [16], and* m_∞ *and* M_∞ *satisfy* (18) *but do not coincide with* m *and* M.

Remark 2 *When the β_k are generated by* (16) *or* (17)*, then, under the conditions of Lemma 1, they asymptotically satisfy $\beta_{2j+1} = M + m - \beta_{2j}$. When ν_{2j} is a two-point measure supported at m and M, we then have $\nu_{2j+2} = \nu_{2j}$. Additionally to Th. 1, $p_1^{(2j)}$ tends to a constant p_∞ as $j \to \infty$, with p_∞ depending on the starting measure ν_0 (*i.e.*, on the starting point x_0) and on the spectrum of A.*

We first consider the case when the α_k used to generate the step-sizes $\gamma_k = 1/\beta_k$ via (15) are i.i.d. with a suitable distribution.

Theorem 2 *Assume that in algorithm* (5) $\beta_k = 1/\gamma_k$ *satisfies* (15) *with $\widehat{m}_k$ and $\widehat{M}_k = \widehat{M}_k^{(i)}$ given by* (14) *for all k for some $i \geq 1$ and that the α_k are i.i.d. in $[-1, 1]$ with a distribution function $F_\alpha(\cdot)$ symmetric with respect to zero. Assume, moreover, that*

$$\int \log(t-u)^2 \, dF_\alpha(t) < \int \log(1-t)^2 \, dF_\alpha(t) < \infty \quad \text{for all } u \in (-1, 1) \tag{20}$$

and that $F_\alpha(1-x) < 1$ for any $x > 0$. Then, $m_\infty = \lim_{k\to\infty} \widehat{m}_k = m$ and $M_\infty = \lim_{k\to\infty} \widehat{M}_k = M$ almost surely.

The proof is given in Section 6. The idea is roughly as follows. In view of Th. 1, the asymptotic behavior of the measures ν_k is very similar to the behavior of measures $\tilde{\nu}_k$ which use the same updating formulas but are supported on the two-point set $\{m, M\}$. However, for this sequence of two-point measures $\tilde{\nu}_k$ the sequence of random variables $\log \tilde{\nu}_k(m) - \log(1 - \tilde{\nu}_k(m))$ is a random walk and therefore the values $\tilde{\nu}_k(m)$ approach 0 and 1 (with any fixed precision) infinitely often. This implies that the sequence of first moments of $\tilde{\nu}_k$ gets arbitrarily close to m and M infinitely often. The same occurs for the original sequence of measures ν_k (proving this requires some technicalities).

As shown in the next theorem, when using $\widehat{M}_k = \widehat{M}_k^{(i)}$ with $i \geq 2$ we do not have to use i.i.d. α_k to obtain $m_\infty = m$ and $M_\infty = M$. Moreover, the sequence $\{\beta_k\}$ can be generated by (16) or (17).

Theorem 3 *Assume that in algorithm* (5) *the $\beta_k = 1/\gamma_k$ are generated according to one of the rules* (15)*,* (16) *or* (17)*, with $\{\alpha_k\}$ having an asymptotic distribution function $F_\alpha(\cdot)$ in $[-1, 1]$ symmetric with respect to zero, and that $\widehat{m}_k$ and $\widehat{M}_k = \widehat{M}_k^{(i)}$ are given by* (14) *for all k for some $i \geq 2$. Assume, moreover, that $F_\alpha(\cdot)$ satisfies* (20) *and that $F_\alpha(1-x) < 1$ for any $x > 0$. Then, $m_\infty = \lim_{k\to\infty} \widehat{m}_k = m$ and $M_\infty = \lim_{k\to\infty} \widehat{M}_k = M$.*

The proof is given in Section 6. The proof of Th. 2 must be modified since, when the α_i are not randomly generated, we cannot be sure to have simultaneously a large value of β_k and a small value of $\mu_1^{(k)}$, see part (ii) of the proof of Th. 2. As a consequence, $i = 1$ does not guarantee the convergence of $\widehat{m}_k$ and $\widehat{M}_k = \widehat{M}_k^{(i)}$ to respectively m and M and we now have to use a more precise estimator of M with $i > 1$.

When the α_k are i.i.d., the result of Th. 3 holds almost surely. The theorem also covers the case where $\{\alpha_k\}$ is a deterministic sequence, for instance generated via a dynamical system, which permits to obtain sequences of rates r_k much less erratic than when using random step-sizes, see Sect. 4. Notice that the proof of Th. 3 does not apply when $\widehat{M}_k = \widehat{M}_k^{(1)} = \max_{j=0,\ldots,k} \mu_1^{(j)}$ (although simulations seem to indicate consistency of $\widehat{m}_k$ and $\widehat{M}_k = \widehat{M}_k^{(1)}$ when (15) or (16) is used, see [16]).

3.3 Controlling the number of updates for $\widehat{m}_k$ and $\widehat{M}_k$

The calculations of $\mu_1^{(k)}$ and $\mu_2^{(k)}/\mu_1^{(k)}$ require the evaluation of several inner products in $\mathbb{R}^n$; therefore, by minimizing the number of iterations where $\widehat{m}_k$ or $\widehat{M}_k$ are updated we can reduce the computational cost of the algorithm. Updates of $\widehat{m}_k$ and $\widehat{M}_k$ in the situation of Th. 2 or Th. 3 can be stimulated by the convergence of the measure ν_k to the set of measures supported at m and M and by the fact that, on the route to this set of two-point measures, ν_k can fluctuate between measures supported at m (when $\widehat{m}_k + \widehat{M}_k > m + M$) and measures supported at M (when $\widehat{m}_k + \widehat{M}_k < m + M$). On the other hand, iterations where $\widehat{m}_k$ (resp. $\widehat{M}_k$) has a good chance to get *significantly* updated are those for which the next measure ν_{k+1} will be close to the delta measure at m (resp. at M). This may happen in particular when β_k is the smallest (resp. largest) among all β_j, $j \leq k$. It can be related to record moments for the α_j and we shall thus consider the situation where updates of $\widehat{m}_k$ or $\widehat{M}_k$ are allowed only at those iterations where α_j is a new record.

For any sequence $\{z_k\} = z_0, z_1, z_2 \ldots$ define the two sequences or record moments $\{L_j^{\min}\} = \{L_j^{\min}\}[\{z_k\}]$ and $\{L_j^{\max}\} = \{L_j^{\max}\}[\{z_k\}]$ by $L_0^{\min} = L_0^{\max} = 0$ and, for all $j \geq 0$,

$$L_{j+1}^{\min} = \min\{k > L_j^{\min} : z_k < z_{L_j^{\min}}\}\,,\ L_{j+1}^{\max} = \min\{k > L_j^{\max} : z_k > z_{L_j^{\max}}\}\,.$$

We also define the numbers of lower and upper record moments, respectively $\delta_j^{\min} = \delta_j^{\min}[\{z_k\}]$ and $\delta_j^{\max} = \delta_j^{\max}[\{z_k\}]$, by

$$\begin{aligned} \delta_j^{\min} &= \#\{i \geq 0 : L_i^{\min} \leq j\} = 1 + \max\{i \geq 0 : L_i^{\min} \leq j\}\,, \\ \delta_j^{\max} &= \#\{i \geq 0 : L_i^{\max} \leq j\} = 1 + \max\{i \geq 0 : L_i^{\max} \leq j\}\,. \end{aligned}$$

For any k, denote by $\bar{\beta}_k$ the value obtained when m and M are substituted for $\widehat{m}_k$ and $\widehat{M}_k$ in (15), (16) or (17). The record moments for $\{\bar{\beta}_k\}$ then coincide with those for $\{\alpha_j\}$, but those for β_k may differ since in general $\widehat{m}_k + \widehat{M}_k \neq m + M$. However, due to Lemma 1, the dissimilarity is asymptotically negligible. When (15) is used, upper (resp. lower) record moments for $\bar{\beta}_k$ coincide with upper (resp. lower) record moments for α_k; when (16) or (17) is used, records for $\bar{\beta}_k$ arrive in pairs, records for $\bar{\beta}_{2j}$ and $\bar{\beta}_{2j+1}$ being associated with a α_j that becomes new record (either lower or upper).

3.3.1 $\{\alpha_k\}$ forms an i.i.d. sequence of random variables

When the α_k are i.i.d., the numbers of lower and upper record moments $\delta_j^{\min}[\{\alpha_k\}]$ and $\delta_j^{\max}[\{\alpha_k\}]$ satisfy $\delta_j^{\min}/\log j \to 1$ and $\delta_j^{\max}/\log j \to 1$ almost surely, see [2, p. 258]. Therefore, when we use (15), $\delta_j^{\min}[\{\bar{\beta}_k\}]/\log j \to 1$ and $\delta_j^{\max}[\{\bar{\beta}_k\}]/\log j \to 1$, whereas $\delta_j^{\min}[\{\bar{\beta}_k\}]/\log j \to 2$ and $\delta_j^{\max}[\{\bar{\beta}_k\}]/\log j \to 2$ when we use (16) or (17).

3.3.2 $\{\alpha_k\}$ is constructed from a low discrepancy sequence

Consider in particular the sequence given by $\alpha_k = \cos(\pi u_k)$ for all $k \geq 0$, with $u_k = (k+1)\varphi \mod 1$ (the fractional part of $(k+1)\varphi$), where $\varphi = (\sqrt{5}+1)/2 \simeq 1.61803\ldots$ is the golden ratio (note that the sequence $\{u_k\}$ can equivalently be constructed through the dynamical system $u_0 = \varphi - 1$, $u_{k+1} = (u_k + \varphi) \mod 1$, $k \geq 0$). This construction is motivated by the associated rate of convergence of the algorithm, see Sect. 4. The corresponding sequences of record moments are $\{L_j^{\min}\}[\{\alpha_k\}] = \{0, 1, 4, 12, 33, \ldots\}$ and $\{L_j^{\max}\}[\{\alpha_k\}] = \{0, 2, 7, 20, 54, \ldots\}$. Denote $\{F_N\}_{N=1}^{\infty} =$

$\{1, 1, 2, 3, 5, 8, 13, 21, 34 \ldots\}$ the sequence of Fibonacci numbers, with exact expression $F_N = (\varphi^N - (-1/\varphi)^N)/\sqrt{5}$. $\{L_j^{\min}\}$ and $\{L_j^{\max}\}$ can then be expressed in terms of the Fibonacci numbers F_j as follows: $L_j^{\min} = F_{2j+1} - 1$, $L_j^{\max} = F_{2j+2} - 1$ for $j = 0, 1, \ldots$ This directly follows from the following two classical results of the theory of Diophantine approximations: (i) for the sequence $\{k\zeta \bmod 1\}$, with any irrational ζ, the successive minimal and maximal values occur when $k = q$ in the denominator of a convergent p/q for ζ in the standard continued fraction expansion of ζ, see [19]; (ii) the convergents of $\zeta = \varphi - 1$ are F_j/F_{j+1} for $j > 1$. The number of upper record moments for $\{\alpha_j\}$ is then $\delta_k^{\max}[\{\alpha_j\}] = 1 + \max\{j : F_{2j+2} - 1 \leq k\} = 1 + \max\{j : F_{2j+2} \leq k + 1\}$. Similar to Proposition 7 in [22] we can show that for all $k > 1$ we have the inequalities $C_0 \log k - 1 < \delta_k < C_0 \log k + 1$ with $C_0 = 1/[2\log(\varphi)] \simeq 1.039$. This yields the asymptotic relation $\delta_k = C_0 \log k + O(1)$ as $k \to \infty$.

When we use (15), the sequence of record moments for $\{\bar{\beta}_k\}$ is the same as for $\{\alpha_k\}$. If we use (17), then the sequences of record moments for $\bar{\beta}_k$ are $\{L_j^{\min}\}[\{\bar{\beta}_k\}] = \{0, 1, 3, 5, 9, 15, \ldots\}$ and $\{L_j^{\max}\}[\{\bar{\beta}_k\}] = \{0, 2, 4, 8, 14, \ldots\}$, with $L_{j+1}^{\min} = L_j^{\max} + 1$ for $j = 0, 1, \ldots$ and, in terms of Fibonacci numbers, $L_j^{\max} = 2(F_{j+2} - 1)$ for $j = 0, 1, \ldots$ The number δ_k of upper record moments for $\{\bar{\beta}_j\}$ thus satisfies $\delta_k / \log k \to 2C_0 = 1/\log(\varphi) \simeq 2.078$ as $k \to \infty$; the same is obviously true for the number of lower record moments. If we use (16), the sequences of record moments for $\bar{\beta}_k$ are $\{L_j^{\min}\}[\{\bar{\beta}_k\}] = \{0, 3, 4, 9, 14, 25 \ldots\}$ and $\{L_j^{\max}\}[\{\bar{\beta}_k\}] = \{0, 1, 2, 5, 8, 15, 24, \ldots\}$; that is, in terms of Fibonacci numbers, $L_j^{\max} = 2(F_{j+1} - 1)$ if j is even and $L_j^{\max} = 2F_{j+1} - 1$ if j is odd, $L_j^{\min} = 2(F_{j+2} - 1)$ if j is even and $L_j^{\max} = 2F_{j+2} - 1$ if j is odd. The numbers of upper and lower record moments satisfy again $\delta_k / \log k \to 1/\log(\varphi) \simeq 2.078$ as $k \to \infty$.

Example 1 We set $n = 800$, $m = 1$ and $M = 1000$, the eigenvalues of A are uniformly distributed in $[m, M]$ and $b = Ac$ in (3, 4) with c uniformly distributed on the unit n-dimensional sphere $\mathscr{S}_n$. We apply the gradient iterations (5) with x_0 uniformly distributed on $\mathscr{S}_n$; the first two iterations correspond to the method of minimum residues with $\gamma_k = \mu_1^{(k)}/\mu_2^{(k)}$, $k = 0, 1$, and the subsequent iterations use $\{\gamma_k\} = \{1/\beta_k\}$, where the β_k are generated via (17) using the low discrepancy sequence $\{\alpha_k\}$ above: $\alpha_k = \cos(\pi[(k+1)\varphi \bmod 1])$ for all $k \geq 0$. We compare the behaviors of two estimators of m and M. The first one corresponds to $\widehat{m}_k$ and $\widehat{M}_k^{(4)}$ given by (14), the second to

$$\widetilde{m}_{2j+1} = \min_{j \in L_\alpha} \mu_1^{(2j+1)} \text{ and } \widetilde{M}_{2j+2} = \max_{j \in L_\alpha} \frac{\mu_4^{(2j+2)}}{\mu_3^{(2j+2)}}\,, \tag{21}$$

where $L_\alpha = \{L_j^{\min}\}[\{\alpha_k\}] \bigcup \{L_j^{\max}\}[\{\alpha_k\}]$ denotes the sequence of lower and upper record moments for $\{\alpha_k\}$. This construction is motivated by the fact that when α_j becomes a new (lower or upper) record, then β_{2j+1} is large and $\mu_1^{(2j+1)}$ has a good chance to be small, while β_{2j+2} is small and $\mu_4^{(2j+2)}/\mu_3^{(2j+2)}$ has a good chance to be large. The precision of the estimation of m and M given by the two estimators is compared in Fig. 1a. Figure 1b indicates the number of record moments for $\{\alpha_k\}$ together with the number of iterations where $\widetilde{m}_k$ and $\widetilde{M}_k$ are updated. One may notice that both $\widetilde{m}_{2j+1}$ and $\widetilde{M}_{2j+2}$ are updated each time α_j is a new record.

4 Fluctuations of the sequence of convergence rates

When ν_k is a two-point measure supported at m and M, applying two successive iterations (11) with $\beta_{k+1} = M + m - \beta_k$ yields $\nu_{k+2} = \nu_k$, with the product of rates $r_k r_{k+1}$ not depending on

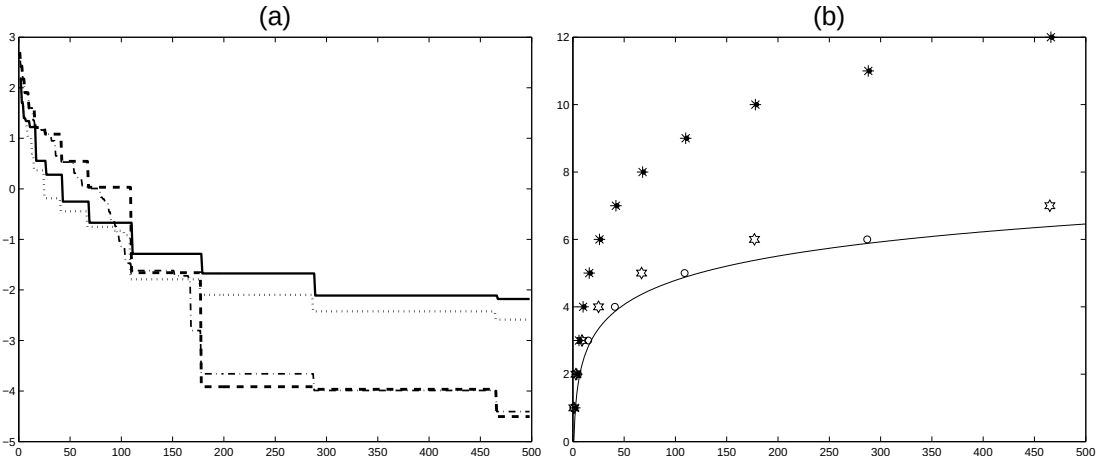

Figure 1: (a) Evolution of $\log_{10}(M-\widehat{M}_k)$ and $\log_{10}(\widehat{m}_k-m)$ with $\widehat{m}_k$ and $\widehat{M}_k=\widehat{M}_k^{(4)}$ given by (14), respectively in dotted and dash-dotted lines, and evolution of $\log_{10}(M-\widetilde{M}_k)$ and $\log_{10}(\widetilde{m}_k-m)$ with $\widetilde{m}_k$ and $\widetilde{M}_k$ given by (21), respectively in solid and dashes lines. (b) Number of record moments $\delta_k^{\max}[\{\alpha_j\}]$ (hexagrams) and $\delta_k^{\min}[\{\alpha_j\}]$ (circles) as functions of k, the number of iterations where $\widetilde{m}_k$ and $\widetilde{M}_k$ are updated are indicated respectively by stars and dots, the solid line corresponds to $(\log k)/[2\log(\varphi)]\simeq 1.039\log k$.

the particular measure ν_k, $r_kr_{k+1}=\mathscr{R}_2^2(\beta_k)$, where

$$\mathscr{R}_2(\beta)=\frac{(M-\beta)(\beta-m)}{\beta(M+m-\beta)}\,.$$

This is the key-point used in [16] to prove the following.

Theorem 4 *Assume that the conditions of Th. 1 are satisfied and that, moreover, the β_k are generated by symmetric pairs for large k; that is, $\beta_{2j+1}=M+m-\beta_{2j}$ for all $j\geq j_0$, with $\beta_{2j}\in[m+\varepsilon,M-\varepsilon]$ for some $\varepsilon\in(0,(M-m)/2)$. Then,*

$$\lim_{k\to\infty}\frac{1}{k}\log R_k=\int\log\left|\frac{(M-t)\,(t-m)}{t\,(m+M-t)}\right|\,dF_\beta(t)=\int\log\frac{(t-m)^2}{t^2}\,dF_\beta(t)\,,\tag{22}$$

where R_k is defined by (6).

Th. 4 applies in particular when $F_\beta(\cdot)$ has the arcsine density $f_\epsilon(\cdot)$ on $[m+\epsilon,M-\epsilon]$,

$$f_\epsilon(\beta)=\frac{1}{\pi\sqrt{(\beta-m-\epsilon)(M-\epsilon-\beta)}}\,,$$

with $\epsilon<(M-m)/2$. In that case, as $k\to\infty$,

$$\begin{aligned}R_k^{1/k}\to R_{\mathrm{arcsine},\epsilon}&=\exp\left\{\int_{m+\epsilon}^{M-\epsilon}\log\frac{(\beta-m)^2}{\beta^2}\,f_\epsilon(\beta)\,d\beta\right\}\\&=\left(\frac{M-m+2\sqrt{\epsilon(M-m-\epsilon)}}{M+m+2\sqrt{(M-\epsilon)(m+\epsilon)}}\right)^2\\&=R_\infty(1+4\sqrt{\epsilon(M-m)})+\mathcal{O}(\epsilon)\,,\ \epsilon\to0\,,\end{aligned}\tag{23}$$

with R_∞ given by (7), see [16]. In the rest of the section we are interested in the extension of Th. 4 to the case where the β_k are generated by (16) or (17) with estimated $\widehat{m}_k$ and $\widehat{M}_k$ and to the fluctuations of $R_k^{1/k}$ along its way to its limiting value.

The fact that in (16, 17) $\widehat{m}_k$ and $\widehat{M}_k$ are estimated brings a slight difference with Th. 4 in terms of asymptotic rate of convergence. This difference is marginal, however, as shown in the next theorem.

Theorem 5 *Assume that in algorithm* (5) *the β_k are generated by pairs as in* (16) *or* (17)*, with $\widehat{m}_k$ and $\widehat{M}_k = \widehat{M}_k^{(i)}$ given by* (14) *for all k for some $i \geq 2$, and that the α_j have an asymptotic distribution function $F_\alpha(\cdot)$ in $[-1+\delta, 1-\delta]$, $0 < \delta < 1/2$, symmetric with respect to zero and satisfying* (20) *and $F_\alpha(1-\delta-x) < 1$ for any $x > 0$. Then the result* (22) *of Th. 4 remains valid, with $F_\beta(\cdot)$ a rescaled version of $F_\alpha(\cdot)$ in some interval $[m+\epsilon', M-\epsilon']$, that is,*

$$dF_\beta(t) = \frac{2}{M-m-2\epsilon'}\, dF_\alpha\left(\frac{2t-m-M}{M-m-2\epsilon'}\right), \tag{24}$$

where ϵ' satisfies

$$\delta(M-m)/2 \leq \epsilon' \leq \sqrt{M}(\sqrt{M}+\sqrt{m})\delta/2 + \mathcal{O}(\delta^2)\,. \tag{25}$$

The proof is given in Sect. 6. The fact that the α_k now lie in $[-1+\delta, 1-\delta]$ with $\delta > 0$ makes it necessary to slightly modify the proof of Th. 3.

Consider now the fluctuations of the asymptotic convergence rate around its limiting value. Suppose that m and M are perfectly estimated, so that $\beta_{2j+1} = M + m - \beta_{2j}$ in (16, 17), with the β_{2j} having a distribution $F_\beta(\cdot)$ satisfying the condition in Th. 4. Since the β_k are exactly symmetric in $[m, M]$, we can take $\delta_1 = 0$ in the proof of Th. 5, so that

$$\log[\mathscr{R}_2(\beta_{2j})] - B\theta^{2j} < \frac{1}{2}\,\log(r_{2j}r_{2j+1}) < \log[\mathscr{R}_2(\beta_{2j})] + B\theta^{2j}$$

for some $B > 0$ and $j > j_1$ large enough, which gives

$$\left|\frac{1}{2}\,\log\left(\frac{R_{2i}}{R_{2j_1}}\right) - \sum_{j=j_1}^{i-1} \log[\mathscr{R}_2(\beta_{2j})]\right| < \sum_{j=0}^{\infty} B\theta^{2j} = \frac{B}{1-\theta}\,,$$

with θ as in Th. 1. Define

$$\begin{aligned}
\mathscr{L}_\alpha &= \int \log[\mathscr{R}_2(t)]\, dF_\beta(t) = \int \log\left(\frac{1+t}{t+\frac{M+m}{M-m}}\right)^2 dF_\alpha(t)\,,\\
\mathscr{V}_\alpha &= \int \left[\log\left(\frac{1+t}{t+\frac{M+m}{M-m}}\right)^2 - \mathscr{L}_\alpha\right]^2 dF_\alpha(t)\,.
\end{aligned}$$

(Note that both quantities are well defined when $F_\alpha(\cdot)$ is concentrated on $[-1+\delta, 1-\delta]$, $\delta > 0$.)

When the α_k are i.i.d. with the distribution $F_\alpha(\cdot)$, we have, for $i \to \infty$,

$$\frac{1}{i}\log\sqrt{R_{2i}} \overset{\text{a.s.}}{\to} \mathscr{L}_\alpha\,, \quad \sqrt{i}\left(\frac{\log\sqrt{R_{2i}}}{i} - \mathscr{L}_\alpha\right) \overset{\text{d}}{\to} \xi \sim \mathscr{N}(0, \mathscr{V}_\alpha)\,,$$

and

$$R_{2i}^{1/2i} \overset{\text{a.s.}}{\to} \exp(\mathscr{L}_\alpha)\,, \quad \sqrt{i}\left(R_{2i}^{1/2i} - \exp(\mathscr{L}_\alpha)\right) \overset{\text{d}}{\to} \xi \sim \mathscr{N}(0, \mathscr{V}_\alpha \exp(2\mathscr{L}_\alpha))\,.$$

Moreover, from the law of the iterated logarithm,

$$\limsup_{i\to\infty} \frac{\log\sqrt{R_{2i}} - i\mathscr{L}_\alpha}{\sqrt{2i\mathscr{V}_\alpha}\sqrt{\log\log i}} = 1 \text{ a.s.} \quad \text{and} \quad \liminf_{i\to\infty} \frac{\log\sqrt{R_{2i}} - i\mathscr{L}_\alpha}{\sqrt{2i\mathscr{V}_\alpha}\sqrt{\log\log i}} = -1 \text{ a.s.}\,,$$

implying that, for any $\varepsilon > 0$,

$$R_{2i}^{1/2i} > \exp\left[\mathscr{L}_\alpha + \frac{(1-\varepsilon)\sqrt{2\mathscr{V}_\alpha \log\log i}}{\sqrt{i}}\right] \tag{26}$$

infinitely often (a.s.), and thus indicating that the fluctuations of the normalized convergence rate $R_{2i}^{1/2i}$ are unavoidably large.

Suppose now that $\{\alpha_k\}$ is constructed from a low-discrepancy sequence, as in Sect. 3.3. Then

$$\left|\frac{1}{i-j_1}\sum_{j=j_1}^{i-1} \log[\mathscr{R}_2(\beta_{2j})] - \mathscr{L}_\alpha\right| < C_\alpha \frac{\log i}{i}$$

for some large enough j_1 (see the proof of Th. 5) and some constant C_α depending on the sequence considered, see, *e.g.*, [10]. Therefore, in that case the normalized convergence rate satisfies $R_{2i}^{1/2i}/\exp(\mathscr{L}_\alpha) < i^{C_\alpha/i}$ and shows much less fluctuations on its route to its limiting value $\exp(\mathscr{L}_\alpha)$ than when the α_k are i.i.d. random variables. Indeed, denoting D_{2i} the difference $R_{2i}^{1/2i} - \exp(\mathscr{L}_\alpha)$, we have

$$\frac{(D_{2i})_{i.i.d.}}{|(D_{2i})_{LDS}|} > \frac{\exp\left[\frac{(1-\varepsilon)\sqrt{2\mathscr{V}_\alpha \log\log i}}{\sqrt{i}}\right] - 1}{i^{C_\alpha/i} - 1} \text{ i.o. a.s. for any } \varepsilon > 0$$

where the right-hand side behaves like $c\sqrt{i\log\log i}/\log i$ as $i \to \infty$ ($c = (1-\varepsilon)\sqrt{2\mathscr{V}_\alpha}/C_\alpha$). This justifies the preference given to low discrepancy sequences over random sequences in the algorithms presented in [22]. One of them is summarized in Sect. 5.

Example 2 We consider the same problem as in Sect. 3.3 (we take now $x_0 = (10^5, 1, 1, \ldots, 1)^\top$ to slow down convergence and better illustrate the different behaviors for the two types of step-size sequences). The β_k are generated by (17) with $\widehat{m}_k$ and $\widehat{M}_k = \widehat{M}_k^{(4)}$ given by (14). Figure 2 shows $R_{2i}^{1/2i}$ as a function of i for the cases when $\{\alpha_k\}$ is the low discrepancy sequence given by $\alpha_k = \cos(\pi[(k+1)\varphi \bmod 1])$ for all $k \geq 0$ and when $\{\alpha_k\}$ is a sequence of i.i.d. random variables having the arcsine distribution; both sequences are generated on $[-1+\delta; 1-\delta]$ with $\delta = 0.005$.

5 Prototype algorithm and simulation results

The estimation of the spectral bounds m and M via (21) permits to construct a gradient algorithm which is quite parsimonious in terms of number of computations of inner products. In order to avoid using multiplications by A when calculating $\mu_1^{(k)}$ and $\mu_4^{(k)}/\mu_3^{(k)}$, the following recursions are used:

$$\mu_1^{(k)} = \frac{(Ag_k, g_k)}{(g_k, g_k)} = \beta_k \left[1 - \frac{(g_k, g_{k+1})}{(g_k, g_k)}\right]$$

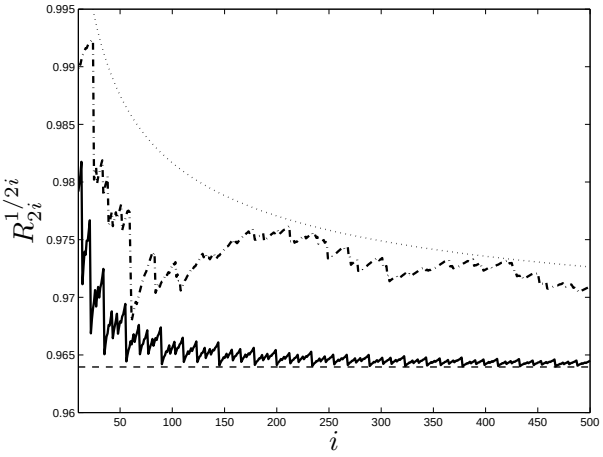

Figure 2: Rate $R_{2i}^{1/2i}$ as a function of i for $\alpha_k = \cos(\pi[(k+1)\varphi \bmod 1])$ for all $k \geq 0$ (solid line) and for $\{\alpha_k\}$ a random i.i.d. sequence having the arcsine distribution (dash-dotted line). The dashed line indicates the bound $R_{\mathrm{arcsine},\epsilon'}$ given by (23) with $\epsilon' = \delta(M-m)/2$ ($\delta = 5 \cdot 10^{-3}$), see (25); the dotted line denotes the right-hand side of (26).

and

$$\frac{\mu_4^{(k-1)}}{\mu_3^{(k-1)}} = \frac{(A^2 g_{k-1}, A^2 g_{k-1})}{(A^2 g_{k-1}, A g_{k-1})} = \beta_{k-1} + \beta_k \frac{(\beta_k(g_{k+1}-g_k) + \beta_{k-1}(g_{k-1}-g_k), g_{k+1}-g_k)}{(\beta_k(g_{k+1}-g_k) + \beta_{k-1}(g_{k-1}-g_k), g_{k-1}-g_k)},$$

which can easily be derived from (2). We generate the $\beta_k = 1/\gamma_k$ according to (17) using the low discrepancy sequence $\alpha_k = \cos(\pi[(k+1)\varphi \bmod 1])$ for all $k \geq 0$. The construction (17) tends to favor the estimation of m against that of M, see [22], which results in the concentration of ν_k at M. From (12), the convergence is monotonic at step k (i.e., $r_k < 1$) when $\beta_k > \mu_2^{(k)}/(2\mu_1^{(k)})$. When ν_k gets close to the delta measure at M, this ratio becomes close to $M/2$ when M/m is large, and the monotonicity condition is violated frequently (approximately every second iteration). In the algorithm proposed below this is avoided by forcing ν_k to become concentrated at m rather than M, the monotonicity property $r_k < 1$ being always satisfied when $\mu_2^{(k)}/(2\mu_1^{(k)})$ is close to $m/2$ since β_k is larger than $\widehat{m}_k > m$. This can be achieved by using a step with large β_k when ν_{k-1} becomes close to the delta measure at M. In practice, we simply use $\beta_k = \widehat{M}_k$ when we observe $\widehat{M}_k > \widehat{M}_{k-1}$. The algorithm is summarized below; its MATLAB implementation is available at `http://www.i3s.unice.fr/~pronzato/Matlab/goldenArcsineQ.m`. We define $v_j = \varphi(j+1) \bmod 1$ and, for $j = 0, 1, \ldots$ we set

$$z_j = (1+\cos(\pi u_j))/2, \text{ where } u_{2j} = \min\{v_j, 1-v_j\},\ u_{2j+1} = \max\{v_j, 1-v_j\}.$$

Algorithm

Stage I (initialization)

I.1 Choose x_0 and compute $g_0 = Ax_0 - b$.

I.2 Choose $\epsilon > 0$ used in the stopping rule.

I.3 Set $L_{\max} = \{2F_{i+2} - 2 : i = 0, 1, \ldots\} = \{0, 2, 4, 8, 14, \ldots\}$.

I.4 For $k = 0$ and 1, set $x_{k+1} = x_k - (1/\beta_k)g_k$ and $g_{k+1} = Ax_{k+1} - b$, where $\beta_k = (Ag_k, Ag_k)/(Ag_k, g_k)$.

I.5 Set $\widehat{m}_2 = \min\{\beta_0, \beta_1\}$ and $\widehat{M}_1 = \widehat{M}_2 = \max\{\beta_0, \beta_1\}$.

I.6 Set $k = 2$ and $j = 0$.

Stage II (iterations)

II.1 If $\widehat{M}_k > \widehat{M}_{k-1}$ then set $\beta_k = \widehat{M}_k$. Otherwise set $\beta_k = \widehat{m}_k + (\widehat{M}_k - \widehat{m}_k) z_j$ and $j \leftarrow j+1$.

II.2 Set $x_{k+1} = x_k - (1/\beta_k) g_k$ and $g_{k+1} = A x_{k+1} - b$.

II.3 If $j-2 \in L_{\max}$ then compute
$\widehat{m}_{k+1} = \min\{\widehat{m}_k, \mu_1^{(k)}\}$, $\widehat{M}_{k+1} = \max\{\widehat{M}_k, \mu_4^{(k-1)}/\mu_3^{(k-1)}\}$,
and check the stopping rule $(g_k, g_k) \leq \epsilon$. Otherwise set $\widehat{m}_{k+1} = \widehat{m}_k$, $\widehat{M}_{k+1} = \widehat{M}_k$.

II.4 Set $k \leftarrow k+1$ and return to Step II.1.

The stopping rule used by the algorithm is simply $(g_k, g_k) < \epsilon$ for some given ϵ. The value of (g_k, g_k) is available for k such that $j - 2 \in L_{\max}$, at such iterations we can thus check the condition $(g_k, g_k) < \epsilon$ directly. Since $\widehat{M}_k/\widehat{m}_k$ provides an under-estimate for ρ, and hence and under-estimate for R_∞ given by (7), we can thus estimate the number of iterations still remaining to achieve the required precision. The proposed algorithm only requires one matrix-vector multiplication per iteration (used to calculate the gradient $g_k = Ax_k - b$), like other gradient methods and Krylov-space based algorithms, in particular CR and CG. When A is sparse, the computation of inner products also contributes significantly to the total computational cost; when using parallel computing with distributed memory machines, it may even yield the main contribution to the efficiency loss, see [21, Sect. 4.4]. The standard formulation of CG (and also CR) requires the computation of two inner products per iteration (some sophisticated versions of CG compute these two inner products in parallel, at the possible expense of a slight increase of storage and maybe reduced numerical stability, see for instance [9, 17]). The prototype algorithm above requires the computation of four inner products in the initial two iterations and then four inner products (possibly computed in parallel) each time the estimates $\widehat{m}_k$ and $\widehat{M}_k$ are updated. This is done when $j - 2 \in L$. Therefore, the total number of inner products computed within $k+1$ steps of the proposed algorithm is equal to $N_k = 4 + 4\delta_j$ where $j = j(k)$ is defined by the algorithm and δ_j satisfies $\delta_j = \log j / \log(\varphi) + O(1)$ as $j \to \infty$. This and the fact that $k/j(k) \to 1$ as $k \to \infty$ imply that the number of inner products computed within $k+1$ steps is approximately $4 + 4 \log k / \log \varphi \simeq 4 + 8.31 \log k$.

Simulation results

Example 3 In this artificial example, A is diagonal with $m = 1$, $M = 1\,000$, $n = 1\,000$ and $b = Ac$ with c random (uniformly distributed on the unit n-dimensional sphere $\mathscr{S}_n$). We consider two configurations for the eigenvalues of A and starting point x_0. In the first case, the n eigenvalues are uniformly distributed in $[m, M]$ and x_0 is random (uniformly distributed on $\mathscr{S}_n$). The second configuration corresponds to the worst-case situation for $n-1$ steps of the Conjugate Residual (CR) algorithm: the eigenvalues are

$$\lambda_i = (M+m)/2 + (M-m)/2 \cos[\pi(i-1)/(n-1)] \text{ for } i = 1, \ldots, n$$

and x_0 is such that the α_i^2 in the decomposition (9) are proportional to $\tau_1^2 = 1/2\lambda_1$, $\tau_j^2 = 1/\lambda_j$ for $j = 2, \ldots, n-1$ and $\tau_n^2 = 1/2\lambda_n$, see [15] for details. Figures 3(a) and 3(b) present the evolution of $\log_{10} \|g_k\|$ as a function of k for the algorithm above (Alg) and the CR algorithm in the two configurations respectively. Since CR is optimal for R_k given by (6), our method is

not competitive in this respect. On the other hand, the evolution of $\log_{10}\|g_k\|$ as a function of the number of inner products computed is plotted in Figures 3(c) and 3(d), showing the gain in complexity of the proposed algorithm compared to CR.

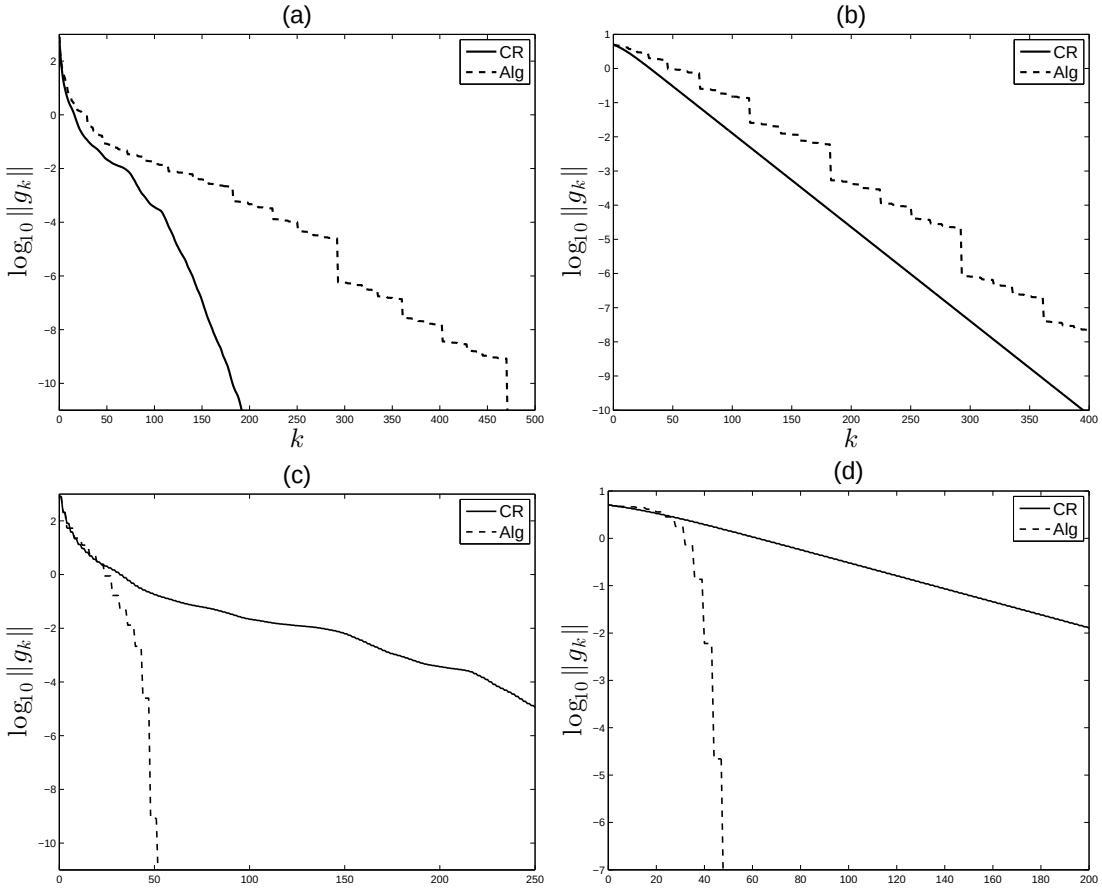

Figure 3: $\log_{10}\|g_k\|$ as a function of k (top), $\log_{10}\|g_k\|$ against the number of inner products computed (bottom) in Example 3. The eigenvalues of A are uniformly distributed in $[m, M] = [1, 1\,000]$ in (a) and (c); (b) and (d) correspond to the worst-case situation for the CR algorithm.

Example 4 A is given by the matrix `NOS5` from `http://math.nist.gov/MatrixMarket/`; it is sparse, with dimension $n = 468$ and $N = 5\,172$ non-zero elements only, its structure is shown in Figure 4 (left). Its condition number approximately equals $1.1 \cdot 10^4$.

Example 5 A is given by the matrix `1138_BUS` from `http://math.nist.gov/MatrixMarket/`; this matrix is also sparse and symmetric positive-definite with $n = 1\,138$ and $N = 4\,054$ non-zero elements, see Figure 4 (right). Its condition number approximately equals $1.5277 \cdot 10^5$.

We set $b = Ac$, c and x_0 are uniformly distributed on $\mathscr{S}_n$. Figure 5 presents the evolution of $\log_{10}\|g_k\|$ against the number inner products in Examples 2 and 3, computed for the algorithm above (Alg) and the CR algorithm. The reduced complexity of the proposed algorithm compared to CR is manifest.

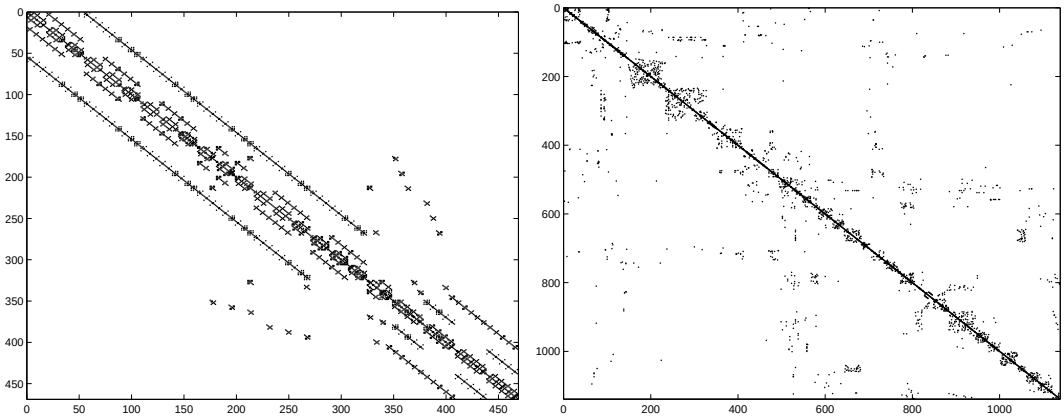

Figure 4: Non-zero elements of A in Examples 4 (left) and 5 (right).

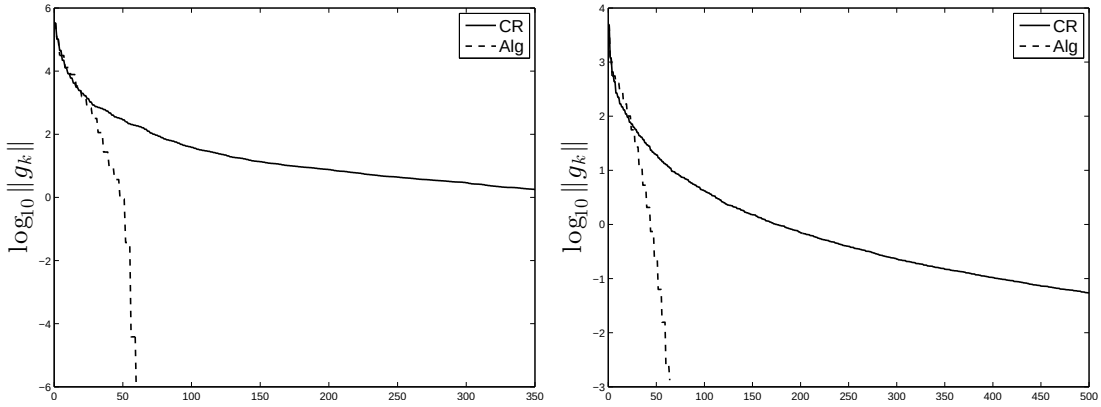

Figure 5: $\log_{10}\|g_k\|$ against the number of inner products computed in Examples 4 (left) and 5 (right).

6 Proofs

Proof of Lemma 1.

Since $\{\widehat{m}_k\}$ forms a non-increasing sequence bounded from below by m, $\widehat{m}_k \to m_\infty$ as $k \to \infty$ for some $m_\infty \geq m$. Similarly, $\widehat{M}_k \to M_\infty$ for some $M_\infty \leq M$. Denote $\epsilon_1 = m_\infty - m$, $\epsilon_2 = M - M_\infty$, $\epsilon_1, \epsilon_2 \geq 0$. Suppose that $0 \leq \epsilon_1 < \epsilon_2$. Then, the asymptotic distribution of the sequence $\{\beta_k\}$ is biased towards m. From (11), the sequence $\{\nu_k\}$ tends to concentrate at M so that $\widehat{M}_k \to M$ as $k \to \infty$, implying $\epsilon_2 = 0$, which contradicts the assumption $0 \leq \epsilon_1 < \epsilon_2$. Similarly, the assumption $\epsilon_1 > \epsilon_2 \geq 0$ leads to a contradiction; therefore, $M - M_\infty = m_\infty - m$. ■

Proof of Th. 2.

From Lemma 1, $m_\infty = m + \epsilon$ and $M_\infty = M - \epsilon$ for some $\epsilon \geq 0$. Assuming that $\epsilon > 0$, we show that this leads to a contradiction. The proof is in two steps. In (i) we show that $\widehat{m}_k$ is repeatedly updated; in (ii) we show that this implies $m_\infty = m$ and thus $\epsilon = 0$.
(i) We have from (11)

$$\frac{p_1^{(k+1)}}{p_d^{(k+1)}} = \frac{(\beta_k - m)^2}{(M - \beta_k)^2} \frac{p_1^{(k)}}{p_d^{(k)}} .$$

When β_k satisfies (15) with $F_\alpha(\cdot)$ satisfying (20) and $\epsilon > 0$, then (19) is satisfied so that $p_d^{(k)} = 1 - p_1^{(k)} - \delta_k$ with

$0 \leq \delta_k \leq D\theta^k$ when $k > k_0$ for some k_0, $D > 0$ and $0 \leq \theta < 1$. Denoting $p_k = p_1^{(k)}$, $Q_k = p_k/(1-p_k)$, we get

$$Q_{k+1} = \frac{p_{k+1}}{(1-p_{k+1})} = Q_k \, \frac{(\beta_k - m)^2}{(M-\beta_k)^2} \, \frac{1 - \frac{\delta_{k+1}}{1-p_{k+1}}}{1 - \frac{\delta_k}{1-p_k}} \, . \tag{27}$$

Since $m + \epsilon \leq \mu_1^{(k)} \leq M - \epsilon$ for all k, we have

$$\frac{\epsilon}{M-m} - B\theta^k \leq p_k \leq 1 - \frac{\epsilon}{M-m} + B\theta^k$$

for some $B > 0$ when $k > k_0$. Therefore, $1-p_k > \epsilon/[2(M-m)]$ for k large enough and, together with $0 \leq \delta_k \leq D\theta^k$, (27) gives

$$\log Q_{k+1} = \log Q_k + \log \left(\frac{\beta_k - m}{M - \beta_k} \right)^2 + c_k \, ,$$

with $|c_k| < C\theta^k$ for all k larger than some k_1, $C = 4D(M-m)/\epsilon$. This implies that

$$\left| \log \frac{Q_{k+1}}{Q_{k_1}} - \sum_{j=k_1}^{k} \log \left(\frac{\beta_j - m}{M - \beta_j} \right)^2 \right| < \sum_{j=0}^{\infty} C\theta^j < \frac{C}{1-\theta} \, . \tag{28}$$

Denote

$$\xi_j = \log \left(\frac{\beta_j - m}{M - \beta_j} \right)^2 .$$

Suppose that there is no update of $\widehat{m}_k$ and $\widehat{M}_k$ after some k_2 and denote $\epsilon_m = \widehat{m}_{k_2} - m$, $\epsilon_M = M - \widehat{M}_{k_2}$. Then, for $j > k_2$, $\{\xi_j\}$ forms a sequence of i.i.d. random variables and (28) indicates that, for $k > k_2$, $\log Q_{k+1} - \log Q_{k_2}$ behaves like a random walk. The random variables ξ_j have mean

$$M(\epsilon_m, \epsilon_M) = \int \log \left(\frac{t + 1 + \frac{2\epsilon_m}{M-m-\epsilon_m-\epsilon_M}}{1 + \frac{2\epsilon_M}{M-m-\epsilon_m-\epsilon_M} - t} \right)^2 .$$

From Lemma 1, $\epsilon_m = \epsilon_M = \epsilon$. Since $M(\epsilon, \epsilon) = 0$, the random walk has no drift and we have $\limsup_{k\to\infty} \log Q_k = -\liminf_{k\to\infty} \log Q_k = \infty$ a.s., which contradicts the assumption of no update of $\widehat{m}_k$ and $\widehat{M}_k$ after iteration k_2. (One may notice that $\xi_{2j} + \xi_{2j+1} = 0$ when the β_k are generated according to (16), so that the argument cannot be used in that case.)

Suppose now that there is no update of $\widehat{m}_k$ for $k > k_2$ and denote $\epsilon_m = \widehat{m}_{k_2} - m$. From the argument above, $\widehat{M}_k$ is repeatedly updated, which, from Lemma 1, is only possible if $\epsilon_{M,k} = M - \widehat{M}_k > \epsilon_m$. The ξ_j, for $j > k_2$, are now neither independent nor identically distributed, but $\mathbb{E}(\xi_j | \mathscr{F}_{j-1}) = M(\epsilon_m, \epsilon_{M,j}) < 0$, with $\{\mathscr{F}_j\}$ the sequence of σ-fields $\sigma(\alpha_0, \alpha_1, \ldots, \alpha_j)$. From (28), the sequence of $\log Q_{k+1} - \log Q_{k_2}$, $k > k_2$, thus forms a supermartingale relative to $\{\mathscr{F}_j\}$. Consider now $S_k = \sum_{j=k_2}^{k} \xi_j - \mathbb{E}(\xi_j | \mathscr{F}_{j-1})$, which forms a martingale sequence. Since we assume that $\epsilon > 0$, the increments $|\xi_j - \mathbb{E}(\xi_j | \mathscr{F}_{j-1})|$ are bounded and $S_k/\sqrt{k}$ satisfies the central limit theorem (see, *e.g.*, [6]). This implies that $\liminf_{k\to\infty} \log Q_k = -\infty$ and therefore $M_\infty = M$, which contradicts the assumption of $\epsilon > 0$. We have thus proved that $\widehat{m}_k$ is updated infinitely often.

(ii) Similarly to (i), $m_\infty = m + \epsilon$ and $M_\infty = M - \epsilon$ with $\epsilon > 0$, β_k satisfying (15) with $F_\alpha(\cdot)$ satisfying (20) imply that (19) is satisfied. Denoting $p_k = p_1^{(k)}$, direct calculations using (11) and (12) give

$$p_{k+1} \geq \left[1 - \frac{1-p_k}{1 + 4p_k\omega_k} \right] + D\theta^k \tag{29}$$

when $k > k_0$, for some constants $D \leq 0$ and $0 \leq \theta < 1$, where $\omega_k = \zeta_k/(1-\zeta_k)^2$ with $\zeta_k = [\beta_k - (M+m)/2]/[(M-m)/2]$ and $\beta_k \in (m+\epsilon, M-\epsilon)$ by construction. (Notice that the term within square brackets on the right-hand side of (29) is an increasing function of both p_k and β_k.) The fact that $F_\alpha(1-x) < 1$ for any $x > 0$ implies that $\limsup_{k\to\infty} \beta_k = M - \epsilon$. We have shown in (i) that $\widehat{m}_k$ is updated infinitely often. Therefore, for any $\delta_1, \delta_2 > 0$, there exists a subsequence $\{j_i\}$ such that $\beta_{j_i} > M - \epsilon - \delta_1$ and $\mu_1^{(j_i)} < m + \epsilon + \delta_2$. For a two-point measure supported at m and M, this second inequality implies $p_{j_i} > (M - \epsilon - m)/(M-m) - \delta_3$, with $\delta_3 \to 0$ as $\delta_2 \to 0$. Due to Th. 1, we thus have

$$p_{j_i} > \frac{M - m - \epsilon}{M - m} - \delta_3 + B\theta^{j_i}$$

for some constant $B \leq 0$ and all $j_i > k_0$. Together with (29), it gives

$$p_{j_i+1} > \frac{(M-m-\epsilon)^3}{(M-m)[(M-m)^2 - 3\epsilon(M-m-\epsilon)]} - \delta_4 \, ,$$

where δ_4 can be made arbitrarily small by taking δ_1, δ_2 small enough and i large enough. This implies that

$$\mu_1^{(j_i+1)} < m + \epsilon + \delta_5 - \frac{\epsilon(M-m-\epsilon)(M-m-2\epsilon)}{(M-m)^2 - 3\epsilon(M-m-\epsilon)} < m + \epsilon$$

for δ_1, δ_2 small enough and i large enough, which contradicts $m_\infty = m + \epsilon$. ■

Proof of Th. 3.

The proof is similar to part (ii) of the proof of Th. 2. From Lemma 1, $m_\infty = m + \epsilon$ and $M_\infty = M - \epsilon$ for some $\epsilon \geq 0$. Suppose that $\epsilon > 0$. When β_k satisfies one of the rules (15-17) with $F_\alpha(\cdot)$ satisfying (20), then (19) is satisfied which implies (29) for $k > k_0$ and some constants $D \leq 0$ and $0 \leq \theta < 1$, where $\omega_k = \zeta_k/(1-\zeta_k)^2$ with $\zeta_k = [\beta_k - (M+m)/2]/[(M-m)/2]$ and $\beta_k \in (m+\epsilon, M-\epsilon)$ by construction. Since $M_\infty = M - \epsilon$, (13) implies that $\mu_2^{(k)}/\mu_1^{(k)} \leq M - \epsilon$. For a two-point measure supported at m and M, this implies $p_k \geq M\epsilon/[(M-m)(m+\epsilon)]$; in view of Th. 1, we thus have

$$p_k \geq \frac{M\epsilon}{(M-m)(m+\epsilon)} + B\theta^k \tag{30}$$

for some constant $B \leq 0$ and $k > k_0$. Since $F_\alpha(1-x) < 1$ for any $x > 0$, $\limsup_{k\to\infty} \beta_k = M - \epsilon$ and, for any $\delta_1 > 0$ and any k_1, there exist some $k > k_1$ such that $\beta_k > M - \epsilon - \delta_1$. In view of (29) and (30) this implies that

$$p_{k+1} \geq \frac{M(M-m-\epsilon)}{(M-m)(M-\epsilon)} - \delta_2$$

where δ_2 can be made arbitrarily small by taking δ_1 small enough and k_1 large enough. This implies in turn that

$$\mu_1^{(k+1)} \leq \frac{Mm}{M-\epsilon} + \delta_3 = m + \epsilon + \delta_3 - \frac{\epsilon(M-m-\epsilon)}{M-\epsilon} < m + \epsilon$$

for δ_1 small enough and k_1 large enough, which contradicts $m_\infty = m + \epsilon$ with $\epsilon > 0$. ■

Proof of Th. 5.

Following the same arguments as in the proof of Th. 3, (30) implies that $\mu_1^{(k+1)} < m + \epsilon$ for $\delta < \delta_\epsilon$ and k large enough, with

$$\delta_\epsilon = \frac{2\epsilon(\sqrt{Mm} - m - \epsilon)(M-m-\epsilon)}{[Mm - (m+\epsilon)^2](M-m-2\epsilon)} = \frac{2\epsilon}{\sqrt{m}(\sqrt{M}+\sqrt{m})} + \mathcal{O}(\epsilon^2)\,.$$

From this we obtain that for small δ, $0 \leq m_\infty - m = M - M_\infty \leq \sqrt{m}(\sqrt{M}+\sqrt{m})\delta/2 + \mathcal{O}(\delta^2)$, so that the β_k are asymptotically distributed in $[m+\epsilon', M-\epsilon']$ with ϵ' satisfying (25). (Note that $m + \epsilon'' \leq \beta_k \leq M - \epsilon''$ for all k, with $\epsilon'' = \delta(M-m)/2$.)

The conditions of Th. 1 are satisfied, so that $\sum_{i=2}^{n-1} \nu_k(\lambda_i) \leq C\theta^k$ for $k > 2j_0$ for some constants $C > 0, j_0 > 0$ and $0 \leq \theta < 1$. Also, accounting for the fact that the distribution of β_{2j} is not exactly symmetric in $[m, M]$, for any $\delta_1 > 0$ there exists some j_1 such that for all $j > j_1$, $|r_{2j}r_{2j+1} - \mathscr{R}_2^2(\beta_{2j})| < \delta_1$. Altogether, for $j > j_1$ large enough,

$$\mathscr{R}_2^2(\beta_{2j}) - D\theta^{2j} - \delta_1 < r_{2j}r_{2j+1} < \mathscr{R}_2^2(\beta_{2j}) + D\theta^{2j} + \delta_1 \tag{31}$$

for some $D > 0$. Since $\beta_{2j}, \beta_{2j+1} \in [m+\epsilon'', M-\epsilon'']$ for all k, $\mathscr{R}_2(\beta_{2j}) > \mathscr{R}_2(m+\epsilon'') = \mathscr{R}_2(M-\epsilon'') = \epsilon''(M+m-\epsilon'')/[(m+\epsilon'')(M-\epsilon'')] > 0$ and

$$2\log[\mathscr{R}_2(\beta_{2j})] - \delta_2 < \log(r_{2j}r_{2j+1}) < 2\log[\mathscr{R}_2(\beta_{2j})] + \delta_2$$

for $j > j_1$, where δ_2 can be made arbitrarily small by taking j_1 large enough. From the definition (6) of R_k, we can thus write

$$\frac{1}{2(i-j_1)}\log\left(\frac{R_{2i}}{R_{2j_1}}\right) = \frac{1}{2(i-j_1)}\sum_{j=j_1}^{i-1}\log(r_{2j}r_{2j+1}) = \frac{1}{(i-j_1)}\sum_{j=j_1}^{i-1}\log[\mathscr{R}_2(\beta_{2j})] + C$$

with $|C| < \delta_2/2$. Therefore,

$$\lim_{k\to\infty}\frac{1}{k}\log R_k = \int \log \frac{(1-\alpha+\frac{2\epsilon'}{M-m-2\epsilon'})(1+\alpha+\frac{2\epsilon'}{M-m-2\epsilon'})}{\left(\frac{M+m}{M-m-2\epsilon'}\right)^2 - \alpha^2}\, dF_\alpha(t)$$

for some ϵ' satisfying (25), which can be written as (22) with $F_\beta(\cdot)$ given by (24). ■

References

[1] J. Barzilai and J.M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8:141–148, 1988.

[2] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events*. Springer, Berlin, 1997.

[3] B. Fischer and L. Reichel. A stable Richardson iteration method for complex linear systems. *Numer. Math.*, 54:225–242, 1988.

[4] G. E. Forsythe. On the asymptotic directions of the s-dimensional optimum gradient method. *Numer. Math.*, 11:57–76, 1968.

[5] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, third edition, 1996.

[6] P. Hall and C.C. Heyde. *Martingale Limit Theory and Its Applications*. Academic Press, New York, 1980.

[7] M. H. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.*, 49:409–436, 1952.

[8] M.A. Krasnosel'skii and S.G. Krein. An iteration process with minimal residues. *Mat. Sb. (in Russian)*, 31(4):315–334, 1952.

[9] G. Meurant. The block preconditioned conjugate gradient method on vector computers. *BIT*, 24:623–633, 1984.

[10] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, 1992.

[11] C.C. Paige and M.A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12(4):617–629, 1975.

[12] O.M. Podvigina and V.A. Zheligovsky. An optimized iterative method for numerical solution of large systems of equations based on the extremal property of zeros of Chebyshev polynomials. *J. of Scientific Computing*, 12(4):433–464, 1976.

[13] L. Pronzato, H.P. Wynn, and A. Zhigljavsky. Renormalised steepest descent in Hilbert space converges to a two-point attractor. *Acta Appl. Math.*, 67:1–18, 2001.

[14] L. Pronzato, H.P. Wynn, and A.A. Zhigljavsky. Asymptotic behaviour of a family of gradient algorithms in $\mathbb{R}^d$ and Hilbert spaces. *Mathematical Programming*, A107:409–438, 2006.

[15] L. Pronzato, H.P. Wynn, and A.A. Zhigljavsky. A dynamical-system analysis of the optimum s-gradient algorithm. In L. Pronzato and A.A. Zhigljavsky, editors, *Optimal Design and Related Areas in Optimization and Statistics*, chapter 3, pages 39–80. Springer, 2009.

[16] L. Pronzato and A. Zhigljavsky. Gradient algorithms for quadratic optimization with fast convergence rates. *Comput. Optim. Appl.*, 50(3):597–617, 2011.

[17] Y. Saad. Practical use of polynomial preconditionings for the conjugate gradient method. *SIAM J. Sci. Stat. Comp.*, 6(4):865–881, 1985.

[18] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, 2008.

[19] B. Slater. Gaps and steps for the sequence $n\,\theta$ mod 1. *Math. Proc. Camb. Phil. Soc.*, 63:1115–1123, 1967.

[20] H. Tal-Ezer. Polynomial approximation of functions of matrices and applications. *J. of Scientific Computing*, 4(1):25–60, 1989.

[21] H.A. van der Vorst. *Iterative Methods for Large Linear Systems*. Utrecht University, Utrecht, 2000.

[22] A. Zhigljavsky, L. Pronzato, and E. Bukina. An asymptotically optimal gradient algorithm for quadratic optimization with low computational cost. *Optimization Letters*, 2012. (DOI 10.1007/s11590-012-0491-7, to appear).